

PATENT

81940.0052

Express Mail Label No. EV 325 216 981 US

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

Isamu KUROKAWA et al.

Serial No: Not assigned

Filed: July 28, 2003

For: INFORMATION PROCESSING SYSTEM

Art Unit: Not assigned

Examiner: Not assigned

TRANSMITTAL OF PRIORITY DOCUMENT

Mail Stop PATENT APPLICATION

Commissioner for Patents

P.O. Box 1450

Alexandria, VA 22313-1450

Dear Sir:

Enclosed herewith is a certified copy of Japanese patent application No. 2002-240719 which was filed August 21, 2002, from which priority is claimed under 35 U.S.C. § 119 and Rule 55.

Acknowledgment of the priority document(s) is respectfully requested to ensure that the subject information appears on the printed patent.

Respectfully submitted,

HOGAN & HARTSON L.L.P.

Date: July 28, 2003

By: 

Anthony J. Orlor

Registration No. 41,232

Attorney for Applicant(s)

500 South Grand Avenue, Suite 1900

Los Angeles, California 90071

Telephone: 213-337-6700

Facsimile: 213-337-6701

日 本 国 特 許 庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日
Date of Application:

2002年 8月21日

出 願 番 号
Application Number:

特願2002-240719

[ST.10/C]:

[JP 2002-240719]

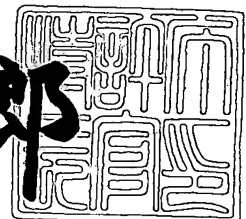
出 願 人
Applicant(s):

株式会社日立製作所

2003年 4月25日

特 許 庁 長 官
Commissioner,
Japan Patent Office

太田 信一郎



出証番号 出証特2003-3030685

【書類名】 特許願

【整理番号】 KN1461

【提出日】 平成14年 8月21日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/06

【発明者】

 【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社 日立
 製作所 RAIDシステム事業部内

 【氏名】 黒川 勇

【発明者】

 【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社 日立
 製作所 RAIDシステム事業部内

 【氏名】 中西 弘晃

【発明者】

 【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社 日立
 製作所 RAIDシステム事業部内

 【氏名】 武藤 淳一

【発明者】

 【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社 日立
 製作所 RAIDシステム事業部内

 【氏名】 竹内 久治

【発明者】

 【住所又は居所】 神奈川県小田原市中里 3 2 2 番地 2 号 株式会社 日立
 製作所 RAIDシステム事業部内

 【氏名】 川口 勝洋

【特許出願人】

 【識別番号】 000005108

 【氏名又は名称】 株式会社 日立製作所

【代理人】

【識別番号】 100093492

【弁理士】

【氏名又は名称】 鈴木 市郎

【電話番号】 03-3591-8550

【選任した代理人】

【識別番号】 100078134

【弁理士】

【氏名又は名称】 武 顕次郎

【電話番号】 03-3591-8550

【手数料の表示】

【予納台帳番号】 113584

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 情報処理システム

【特許請求の範囲】

【請求項 1】 上位処理装置と、上位処理装置が入出力対象とするデータを 1 台以上の物理デバイスを用いて記憶する外部記憶装置と、前記上位処理装置と前記外部記憶装置との間に介在してデータ授受の制御を行う制御装置と、を備えた情報処理システムにおいて、

前記制御装置は、前記データを或る大きさを分割管理されたブロックを一時的に保持するキャッシュメモリと、前記外部記憶装置と前記キャッシュメモリの使用状況をモニタリングした情報を記憶する制御情報用メモリと、前記ブロックの集合を論理デバイスとして管理するとともに前記キャッシュメモリと前記制御情報用メモリに接続されて前記ブロックの入出力制御を行う制御部と、を有し、

前記制御装置は、前記上位処理装置により入出力要求された前記ブロックを前記上位処理装置へ転送する際に、現時点で発生している前記入出力要求が限定個数のデータ空間群の中の特定データ空間であるか否かを判断し、

更に、現時点で行う先読み処理が他の入出力要求に与える影響度の大小を判断し、

現時点で発生している入出力要求が前記特定データ空間であり且つ現時点で行う先読み処理が他の入出力要求に与える影響度が小さいと判断した場合に、前記上位処理装置が入出力要求する前記ブロックとは別に前記特定データ空間内の 1 つ以上のブロックを前記外部記憶装置から前記キャッシュメモリへロードする

ことを特徴とする情報処理システム。

【請求項 2】 請求項 1 の情報処理システムにおいて、

現時点で発生している入出力要求が前記特定データ空間であると判断するに際して、前記入出力要求が前記データ空間群に在るか否かを判断し、在れば I / O カウントをインクリメントし、無ければ最小の I / O カウントのデータ空間と入れ替えし、I / O カウントが一定値を超える場合に前記特定データ空間であると判断する

ことを特徴とする情報処理システム。

【請求項3】 請求項1又は2の情報処理システムにおいて、

現時点で行う先読み処理が他の入出力要求に与える影響度の大小を判断するに際して、前記データ空間群の中でI/Oカウントが一定値を超える特定データ空間のそれぞれの容量の総和が、前記キャッシュメモリの使用可能領域容量以下である場合に前記影響度が小さいと判断する

ことを特徴とする情報処理システム。

【請求項4】 請求項1、2又は3の情報処理システムにおいて、

前記データ空間は、一の論理デバイス又は前記一の論理デバイス上の或る特定領域を対象とすることを特徴とする情報処理システム。

【請求項5】 請求項1、2又は3の情報処理システムにおいて、

前記データ空間は、論理グループ化された外部記憶装置を対象とすることを特徴とする情報処理システム。

【請求項6】 請求項1の情報処理システムにおいて、

前記外部記憶装置から前記キャッシュメモリへロードする際に、前記キャッシュメモリの使用可能領域容量を越えないように前記キャッシュメモリへのロードする前記ブロック数を変動させることを特徴とする情報処理システム。

【請求項7】 請求項1の情報処理システムにおいて、

現時点で発生している入出力要求が前記特定データ空間でない、又は現時点で行う先読み処理が他の入出力要求に与える影響度が大きいと判断した場合に、前記上位処理装置が入出力要求する前記ブロック以外のブロックを前記外部記憶装置から前記キャッシュメモリへロードしない

ことを特徴とする情報処理システム。

【請求項8】 請求項1乃至7のいずれか1つの請求項の情報処理システムにおいて、

前記ブロックは、論理トラックを対象とすることを特徴とする情報処理システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、上位処理装置と、外部記憶装置と、上位処理装置と外部記憶装置の間に介在する制御装置と、を含めた情報処理システムに関し、特に、上位処理装置からのランダムアクセス入力要求処理における外部記憶装置からの制御装置内キャッシュメモリへのデータ先読み制御技術に関する。

【 0 0 0 2 】

【従来技術】

外部記憶装置からの制御装置内キャッシュメモリへの先読み制御技術においては、上位処理装置からのアクセスパターンの履歴情報に基づき、未来のアクセスパターンを予測し、先読みを制御するものが一般的である。この予測においては、アクセスパターンの履歴情報の蓄積が重要となる。

【 0 0 0 3 】

このようなアクセスパターン予測による先読み制御に関する技術は、特開平 6 - 1 1 9 2 4 4 号公報に開示されている。この公報の技術は、上位処理装置から最近受け付けた複数のアクセス要求の際に、キャッシュヒット／ミスの情報を、アクセスを受けた個々の物理的な領域毎に統計情報として管理記憶し、その後、同一の領域をアクセスする際に、統計情報を参照して、キャッシュにロードしておく場合のキャッシュヒットの可能性を判定し、可能性が高い場合、近傍の領域を含めキャッシュメモリにロードするという制御を行うものである。

【 0 0 0 4 】

また、キャッシュメモリへロードする量の最適化を行う技術は、特開平 8 - 1 6 4 6 6 号公報に開示されている。この公報の技術は、アクセス対象ブロックの過去のキャッシュヒット／ミス回数からヒット率を算出し、ヒット率の水準に応じてキャッシュメモリへロードするデータ量を動的に変動させるというものである。

【 0 0 0 5 】

【発明が解決しようとする課題】

データベースを一定期間に集中してランダムアクセスするような処理が発生した場合に、アクセスパターンの履歴情報などから学習し、キャッシュへのローディングモードを決定する従来技術の方式では、ランダム性のために感度が鈍く最

適なローディングモードとはなり得ず、急激なキャッシュヒット率の向上は実現できず、上位処理装置からの入出力要求に対してはキャッシュミスとなって物理デバイス（一般には磁気ディスク装置のHDD）へのアクセスが発生し、レスポンスはHDD 1 台の性能（シーク＋ディスク 1 回転待ち＋転送時間）に依存してしまう。

【0006】

したがって、本発明の目的は、ランダムアクセスであっても即効性のある方式で高キャッシュヒット率を実現してレスポンスを向上させるデータ先読み制御方法を提供するにある。

【0007】

【課題を解決するための手段】

前記課題を解決するために、本発明は主として次のような構成を採用する。

上位処理装置と、上位処理装置が入出力対象とするデータを 1 台以上の物理デバイスを用いて記憶する外部記憶装置と、前記上位処理装置と前記外部記憶装置との間に介在してデータ授受の制御を行う制御装置と、を備えた情報処理システムにおいて、

前記制御装置は、前記データを或る大きさに分割管理されたブロックを一時的に保持するキャッシュメモリと、前記外部記憶装置と前記キャッシュメモリの使用状況をモニタリングした情報を記憶する制御情報用メモリと、前記ブロックの集合を論理デバイスとして管理するとともに前記キャッシュメモリと前記制御情報用メモリに接続されて前記ブロックの入出力制御を行う制御部と、を有し、

前記制御装置は、前記上位処理装置により入出力要求された前記ブロックを前記上位処理装置へ転送する際に、現時点で発生している前記入出力要求が限定個数のデータ空間群の中の特定データ空間であるか否かを判断し、

更に、現時点で行う先読み処理が他の入出力要求に与える影響度の大小を判断し、

現時点で発生している入出力要求が前記特定データ空間であり且つ現時点で行う先読み処理が他の入出力要求に与える影響度が小さいと判断した場合に、前記上位処理装置が入出力要求する前記ブロックとは別に前記特定データ空間内の 1

つ以上のブロックを前記外部記憶装置から前記キャッシュメモリへロードする情報処理システム。

【0008】

また、前記情報処理システムにおいて、

現時点で発生している入出力要求が前記特定データ空間であると判断するに際して、前記入出力要求が前記データ空間群に在るか否かを判断し、在ればI/Oカウントをインクリメントし、無ければ最小のI/Oカウントのデータ空間と入れ替えし、I/Oカウントが一定値を超える場合に前記特定データ空間であると判断する情報処理システム。

【0009】

また、前記情報処理システムにおいて、

現時点で行う先読み処理が他の入出力要求に与える影響度の大小を判断するに際して、前記データ空間群の中でI/Oカウントが一定値を超える特定データ空間のそれぞれの容量の総和が、前記キャッシュメモリの使用可能領域容量以下である場合に前記影響度が小さいと判断する情報処理システム。

【0010】

【発明の実施の形態】

本発明の実施形態に係る情報処理システムについて図面を参照しながら以下詳細に説明する。図1は本発明の実施形態に係る、上位処理装置と外部記憶装置と制御装置とを備えた情報処理システムの全体構成を示す図であり、図2は本実施形態に関する外部記憶装置の構成を示す図であり、図3は本実施形態に関する制御装置内の制御情報用メモリに格納される情報内容を示す図である。また、図4は入出力要求が特定のデータ空間（具体的には、データベース）であるか否かを判断する処理フローを示す図であり、図5は先読み処理が他の入出力要求に影響を及ぼさないか否かを判断する処理フローを示す図であり、図6は特定のデータ空間（具体的には、データベース）内で1つ以上のブロック（具体的には、論理トラック）をキャッシュメモリへロードする処理フローを示す図である。また、図7は物理デバイスを論理トラックに基づいて使用する使用態様を示す図である。

【0011】

「情報処理システムの全体構成の説明」

図1に示す情報処理システムは、情報の演算等を行う処理装置10と、データ転送等の入出力制御を行うチャネル11と、外部記憶装置13と、処理装置10及び外部記憶装置13とデータ授受の制御を行う制御装置12と、から構成されている。また、制御装置12は、制御を実施する際に使用する制御情報を記憶する制御情報用メモリ122に対する記録再生機能を有する。この情報処理システムは、処理装置10の命令によりチャネル11、制御装置12内のキャッシュメモリ124を介して、外部記憶装置13内に定義される論理デバイス132に対してデータの入出力命令を行うものである。処理装置10から入出力要求のあったデータは、通常、図2に示す論理トラック20毎に、キャッシュメモリ124へ記憶される。

【0012】

また、制御部121が、外部記憶装置13、物理デバイス131及びキャッシュメモリ124の使用状況をモニタリングした情報についても、制御情報用メモリ122内に記憶される。

【0013】

図2に示す外部記憶装置13は、1台以上の物理デバイス131（例えば、HDD）で構成される。また、処理装置10の入出力対象となるデータは論理デバイス132単位で管理される。論理デバイス132は制御装置12により、1台以上の物理デバイス131のグループ上に、1論理デバイスを制御上の論理範囲で分割した形態で定義され、それぞれの分割された領域が各物理デバイス（例えば、個々のHDD）に分散している。

【0014】

一般に、1論理デバイスを制御上の論理範囲で分割された領域は論理トラック20で定義される。この論理トラック20は、1論理デバイスを制御上の論理範囲で分割した最少単位である論理レコード21で構成される。処理装置10は、入出力対象とするデータが存在する論理デバイス132内の論理トラックアドレスと論理レコードアドレスを指定し、制御装置12に対しデータの入出力命令を

行う。

【 0 0 1 5 】

図 3 に示すアクセス中論理デバイス情報テーブル 1 2 3 (現時点まででアクセスが頻繁に実施されている論理デバイスのテーブル) は、番号 0 ~ X までの (X + 1) 個の要素で構成され、それぞれの構成内容は、論理デバイスアドレスと I / O カウントである。具体的に云えば、テーブルの個々のデバイスアドレスは、図 1 の例で論理デバイス A か B か C のいずれか 1 つの論理デバイスの特定であり、I / O カウントは I / O 要求の回数であり、番号 0 ~ X は、現時点までにアドレスした論理デバイスの数を特定したものであって、具体的には一例として、論理デバイスの数が数百 ~ 数千のオーダーである場合に X が数個 ~ 数十個であるというオーダーである。この (X + 1) 個がキャッシュメモリへの先読みデータの対象となる論理デバイスの数を表す。

【 0 0 1 6 】

(X + 1) 個は任意の値ではあるが、特定のデータ空間を識別するために上限値は存在する。図 3 に示す論理デバイスのリストは、キャッシュメモリの容量やアクセスの頻度を勘案してそのリスト個数を決めるものであり、このリストは先読みデータをキャッシュメモリに転送する論理デバイスの候補群である。ここで、データ空間は、上位装置が或る処理を実施するためにアクセスするデータの大きさを云い、アクセスされる論理領域におけるデータ領域である。具体的一例としては、取り扱うデータベースのことを指すが、これに限らず、論理デバイス、論理デバイス上のある特定領域、RAID 上の外部記憶装置であってもよい。

【 0 0 1 7 】

また、上位処理装置が入出力要求に先立ちアクセス範囲を論理トラックの先頭と終端アドレスにて指定する EXTENT 情報を、アクセス中論理デバイス情報テーブル 1 2 3 の要素内の構成内容に追加することで、論理デバイスの特定の領域への入出力要求までの分解能を持つことも可能である。

【 0 0 1 8 】

「入出力要求が特定のデータ空間 (データベース) であることを判断する処理に関する説明」

入出力要求が特定のデータ空間であることを判断する処理をフローチャートで示したのが、図4である。ステップ401では、処理装置10は、入出力対象とする論理デバイスアドレスA、データが存在する論理デバイス132内の論理トラックアドレス α と論理レコードアドレスを指定し、制御装置12に対しデータの入出力命令を行う。処理装置10の命令により、制御装置12は、論理トラックアドレス α に対し入出力処理を実施する。

【0019】

ステップ402では、アクセス中論理デバイス情報テーブル123内に入出力対象とする論理デバイスアドレスAと一致するものが存在するかを検索する。一致するものがある場合は、ステップ403へ続く。ステップ403では、アクセス中論理デバイステーブル123内の論理デバイスアドレスAのI/Oカウントをインクリメントする。

【0020】

一方、一致するものが無い場合、ステップ404では、アクセス中論理デバイス情報テーブル123内の更新ポインタで示される該当情報（直近にアクセスされた論理デバイスについての情報）を除いた情報内で、I/Oカウントが最小であるアクセス中論理デバイスアドレス情報にデバイスアドレスとしてAを記憶する。すなわち、処理装置10からの入出力要求の対象となった論理デバイスが図3に示す情報テーブル123の0～Xに存在しない場合には、直近にアクセスされた論理デバイスを除いて、存在する0～X中でI/Oカウントの最小のものと入れ替えて図3中のリストに加えるのである。

【0021】

ステップ405では、論理デバイスアドレスAを記憶した位置を更新ポインタへ記憶する（新たに図3のリストに加わった論理アドレスAには、直近のアクセスがあったことを記憶しておく）。ステップ406では、論理デバイスアドレスAのI/Oカウントを初期化する。

【0022】

ステップ407では、入出力対象である論理デバイスアドレスのI/Oカウントが一定値より大きいかを判断する。判断結果が真であるならば、入出力要求

が特定のデータ空間（データベース）をアクセスしていると認知する。一方、判断結果が、偽である場合には入出力要求が特定のデータ空間をアクセスしていないと認知する。

【 0 0 2 3 】

以上のように、入出力要求が図 3 に示すテーブル中の論理デバイスであれば I / O カウントを + 1 し、図 3 のリスト外であればリスト入れ替えを行う。そして、入出力要求の最新のアドレスには更新ポインタを付す。さらに、I / O カウントが一定以上になった論理デバイスが特定のデータ空間にアクセスしたものと判断されて、この特定のデータ空間である論理デバイスのデータが先読みされ得る対象とされるのである。

【 0 0 2 4 】

「現時点で行う先読み処理が他の入出力要求に与える影響度を認知する手段に関する説明」

現時点で行う先読み処理が他の入出力要求に与える影響度を認知する処理をフローチャートで示したのが、図 5 である。

【 0 0 2 5 】

ステップ 5 0 1 からステップ 5 0 4 においては、アクセス中論理デバイス情報テーブル内の I / O カウントが一定値を越える論理デバイスそれぞれの容量の総和からアクセス中論理デバイス総容量を算出する。具体的には、図 3 に示すテーブルの X リストの中で I / O カウントが一定値を超えるもののそれぞれの論理デバイスの容量の総和であり、アドレス N を最初は 0 として順次にアドレスをインクリメントして最後に X としてループを形成して求める手順を示している。

【 0 0 2 6 】

ステップ 5 0 5 では、制御部 1 2 1 がモニタリングしているキャッシュメモリ使用状況からキャッシュメモリ使用可能領域容量（キャッシュ容量－ライトデータ使用量）を算出する。ステップ 5 0 6 では、キャッシュメモリ使用可能領域容量よりもアクセス中論理デバイス総容量が小さいかを判断する。判断結果が真であるならば、他の入出力要求に影響を及ぼさないと認知する。一方、判断結果が、偽である場合に他の入出力要求に影響を与えると認知する。

【 0 0 2 7 】

「特定のデータ空間内における1つ以上のブロックをキャッシュメモリへロードする手段に関する説明」

特定のデータ空間（データベース）内における1つ以上のブロック（論理トラック）をキャッシュメモリへロードする処理をフローチャートで示したのが、図6である。

【 0 0 2 8 】

ステップ601と602では、入出力要求が特定のデータ空間をアクセスしていること（図4の処理フローを参照）と、他の入出力要求に影響を及ぼさないこと（図5の処理フローを参照）の両方が成立していることを判断する。判断結果が、偽である場合、入出力要求対象である論理トラックアドレス α 以外のトラックをキャッシュメモリへロードすることは実施しない。

【 0 0 2 9 】

一方、判断結果が真であるならば、ステップ603では、キャッシュメモリへのロード開始位置を入出力要求対象である論理トラックアドレス $(\alpha + 1)$ と設定する。すなわち、入出力要求の対象であった論理トラックアドレス α の次の論理トラックアドレスである $(\alpha + 1)$ を先読みしてキャッシュメモリにロードしようとするものである。ステップ604では、キャッシュメモリ124へのロード数を一定値と決定する（先読みする論理トラックの数を一定値とする）。

【 0 0 3 0 】

ステップ605では、図5中ステップ505で算出したキャッシュメモリ使用可能領域容量よりも一定値でロードするトラック数の総容量（先読みのデータ容量）が小さいかを判断する。判断結果が、偽である場合、ステップ607では、キャッシュメモリ124へのロード数を最小値とする（例えば、ロード数を1とする）。なお、外部記憶装置からキャッシュメモリへロードする際に、キャッシュメモリの使用可能領域容量を越えないようにキャッシュメモリへのロードするブロック数（例えば、論理トラック数）を動的に変動させるようにしても良い。一方、判断結果が真であるならば、ステップ606では、制御部121がモニタリングしている外部記憶装置13の使用状況から使用率が極めて低いかを判断す

る。

【 0 0 3 1 】

判断結果が、偽である場合、キャッシュメモリ 1 2 4 へのロード数は一定値のままとする。一方、判断結果が真であるならば、ステップ 6 0 8 では、制御部 1 2 1 がモニタリングしている外部記憶装置 1 3 の使用状況から使用率がゼロであるかを判断する。判断結果が真であるならば、ステップ 6 0 9 では、キャッシュメモリ使用可能領域容量よりも該当の外部記憶装置 1 3 の総容量が小さいかを判断する。判断結果が真であるならば、ステップ 6 1 0 では、キャッシュメモリ 1 2 4 へのロード数を該当の外部記憶装置内の全トラック数とする。この時、ステップ 6 0 3 で決定したロード開始位置は無効となる。

【 0 0 3 2 】

一方、ステップ 6 0 8 での判断結果が偽である場合（該当の外部記憶装置の使用率が極めて低い（0 ではない）場合）、もしくは、ステップ 6 0 9 での判断結果が偽である場合（キャッシュ使用可能領域容量が該当外部記憶装置の総容量より大でない場合）、ステップ 6 1 1 では、キャッシュメモリ使用可能領域容量よりも、上位処理装置 1 0 が入出力要求に先立ちアクセス範囲を論理トラックの先頭と終端アドレスにて指定する E X T E N T 情報に対応するトラック数の総容量が小さいかを判断する。判断結果が真であるならば、ステップ 6 1 2 では、キャッシュメモリ 1 2 4 へのロード数を E X T E N T 情報に対応するトラック数とする。この時も、ステップ 6 0 3 で決定したロード開始位置は無効となる。一方、判断結果が偽である場合、キャッシュメモリ 1 2 4 へのロード数は一定値のままとする。

【 0 0 3 3 】

以上説明した図 4、図 5 及び図 6 に示す処理フローを用いて、外部記憶装置 1 3 1 内の物理デバイスより、キャッシュメモリ 1 2 4 へデータをロードする。この時の物理デバイス 1 3 1 の使用状態について示したのが図 7 である。図 7 に示す符号 7 0 1 ~ 7 0 3 は、それぞれの物理デバイス 1 3 1 に論理デバイス 1 3 2 の論理トラック 2 0 が配置されている状態を示すものである。

【 0 0 3 4 】

図 7 の符号 7 0 1 ~ 7 0 3 での各樹目は各物理デバイス 1 3 1 内での連続するアドレス空間である。一方、図 7 の符号 7 0 1 ~ 7 0 3 の各樹目上に配置されている論理トラック 2 0 は、当該物理デバイス 1 3 1 内で隣接する樹目内の論理トラック 2 0 との関係において、必ずしも論理デバイス 1 3 2 内での連続するアドレス空間に対応しているわけではない。

【 0 0 3 5 】

隣接する物理デバイス 1 3 1 内の隣接する論理トラック 2 0 との関係において、論理デバイス 1 3 2 内での連続するアドレス空間に対応している場合もある。さらには、図 7 の 7 0 1 ~ 7 0 3 での縦横に隣接する各樹目上の論理トラック 2 0 同士には、論理デバイス 1 3 2 内での連続するアドレス空間に対応していない場合もある。

【 0 0 3 6 】

当該物理デバイス 1 3 1 内で隣接する樹目内の論理トラック 2 0 との関係において、論理デバイス 1 3 2 内での連続するアドレス空間に対応しているのであれば、特定の物理デバイス 1 3 1 の縦方向にロード処理が発生する。隣接する物理デバイス 1 3 1 内の隣接する論理トラック 2 0 との関係において、論理デバイス 1 3 2 内での連続するアドレス空間に対応しているのであれば、横方向に処理が必要であり、複数の物理デバイス 1 3 1 でロード処理が発生する。

【 0 0 3 7 】

以上のように動作する物理デバイス 1 3 1 の数に差ができるのみであり、本発明は、いかなる論理トラック 2 0 の配置状態であってもその実施において影響を受けるものではない。

【 0 0 3 8 】

以上説明したように、本発明においては、従来技術のように今後のアクセスパターンを予測するのではなく、現在の入出力要求に対して、キャッシュメモリをある程度占有でき他の入出力要求に影響を与えることが無いと判断した場合に、一回の HDD へのアクセスで、I/O 要求されたブロック（データを分割した単位であり、具体的には単一の論理トラックであるが、2 つの論理トラックをまとめて 1 つのブロックとしても良い）よりも十分大きな複数ブロックを含んでキャ

ッシュメモリへロードする手段を設ける。

【0039】

すなわち、予測するのではなく、とりあえず読んでおくということで、現在発生している入出力要求に対し、制御装置配下の資源（キャッシュメモリやバスなど）の使用率を最大限にするということが目的である。ランダムアクセスといっても、データベースのサイズが有限である以上は、以前アクセスした領域の近傍にアクセスがある可能性は高い。

【0040】

制御装置に搭載するキャッシュメモリが増大化している昨今では、この手段にて、後続のランダムI/OはキャッシュヒットできHDDへのアクセスが無くなるのでHDD1台の性能以上のレスポンスを得ることができる。

【0041】

特に、RAIDでは複数台のHDDからRAIDグループを構成し、RAIDグループを構成するHDDは大容量なので、このRAIDグループに複数のデータベースが格納される。ランダムアクセスであることを考慮すれば、次のI/O要求は同一RAIDグループ内の別HDDへのアクセスである確率が高いので、前回のI/O要求で1台のHDDを占有していても、次のI/Oとの衝突（或る1つのHDDに対して、前回のI/O処理継続中（転送速度が遅い）に次のI/O要求が発生すること）は少ないので、別HDDへのランダムアクセスであるRAIDでは特に効果的である。逆に、RAIDグループ内の複数データベースへのアクセスが発生している場合では、1回のI/Oで長時間ある1台のHDDを占有することは、他のI/Oとのぶつかりによる待ちでレスポンス低下になるので、多数のデータベースへのアクセスが発生していないことを判断する手段を設ける必要がある。

【0042】

このように、本発明の実施形態においては、従来技術に示すような、キャッシュヒット又はミス情報をデータ領域毎に記憶管理して、このアクセスパターンの履歴情報に基づいてキャッシュメモリにロードするという、過去のアクセスパターンを利用するものではなくて、最新の入出力要求のあったデータベースを重

視し、キャッシュメモリの使用可能容量を最大限活用して、入出力要求のあった論理トラックに続く論理トラックを先読みロードするものであり、その際に I / O カウントが一定値を超えるデータベースであることを勘案するので、最新の且つ I / O 頻度の多いデータベースの先読みを行って、キャッシュヒットの観点並びにメモリの利用効率の観点で効率性を高めたものである。

【 0 0 4 3 】

【発明の効果】

データベースを一定期間に集中してランダムアクセスするような処理が発生した場合に、アクセスパターンの履歴情報などから学習し、キャッシュへのローディングモードを決定する従来の発明方式ではなく、制御装置内資源の使用率を最大限引き出す処理を実施することで即効性を高め、ランダムアクセスであっても高キャッシュヒット率を実現しレスポンスを向上することができる。

【図面の簡単な説明】

【図 1】

本発明の実施形態に係る、上位処理装置と外部記憶装置と制御装置とを備えた情報処理システムの全体構成を示す図である。

【図 2】

本実施形態に関する外部記憶装置の構成を示す図である。

【図 3】

本実施形態に関する制御装置内の制御情報用メモリに格納される情報内容を示す図である。

【図 4】

入出力要求が特定のデータ空間（データベース）であるか否かを判断する処理フローを示す図である。

【図 5】

先読み処理が他の入出力要求に影響を及ぼさないか否かを判断する処理フローを示す図である。

【図 6】

特定のデータ空間（データベース）内で 1 つ以上のブロック（論理トラック）

をキャッシュメモリへロードする処理フローを示す図である。

【図 7】

物理デバイスを論理トラックに基づいて使用する使用態様を示す図である。

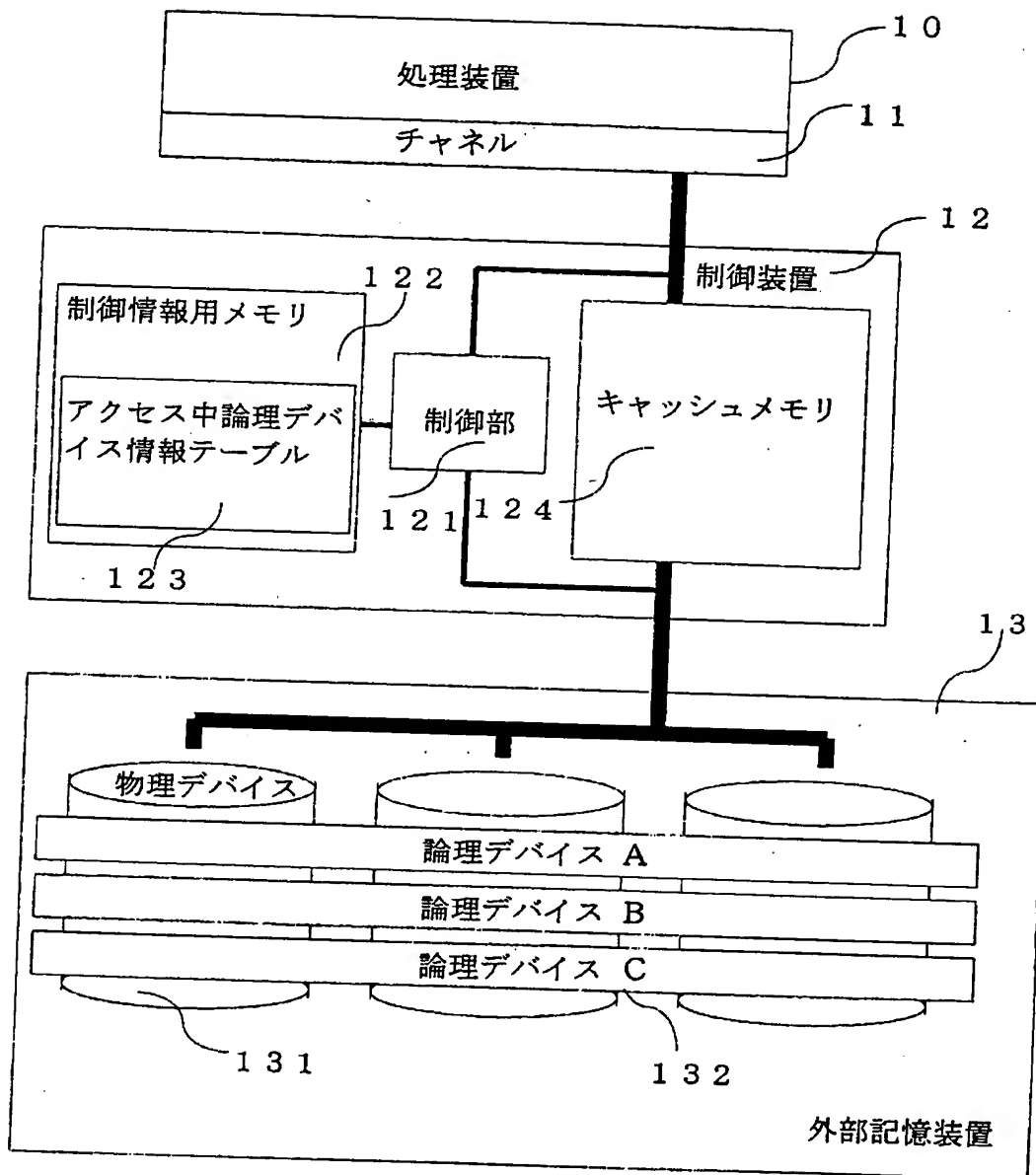
【符号の説明】

- 1 0 処理装置
- 1 1 チャンネル
- 1 2 制御装置
- 1 3 外部記憶装置
- 1 2 1 制御部
- 1 2 2 制御情報用メモリ
- 1 2 3 アクセス中論理デバイス情報テーブル
- 1 2 4 キャッシュメモリ
- 1 3 1 物理デバイス
- 1 3 2 論理デバイス
- 2 0 論理トラック
- 2 1 レコード
- 7 0 1, 7 0 2, 7 0 3 論理トラック配置状態

【書類名】 図面

【図 1】

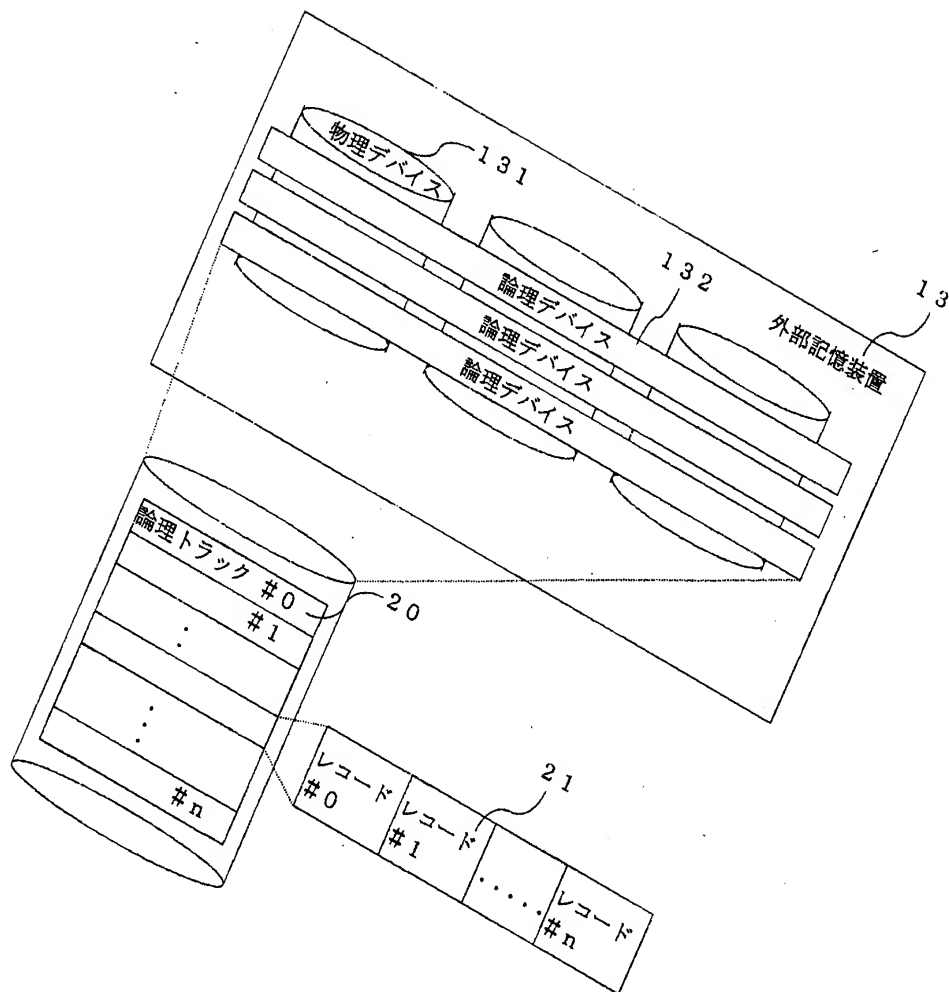
図 1



【図2】

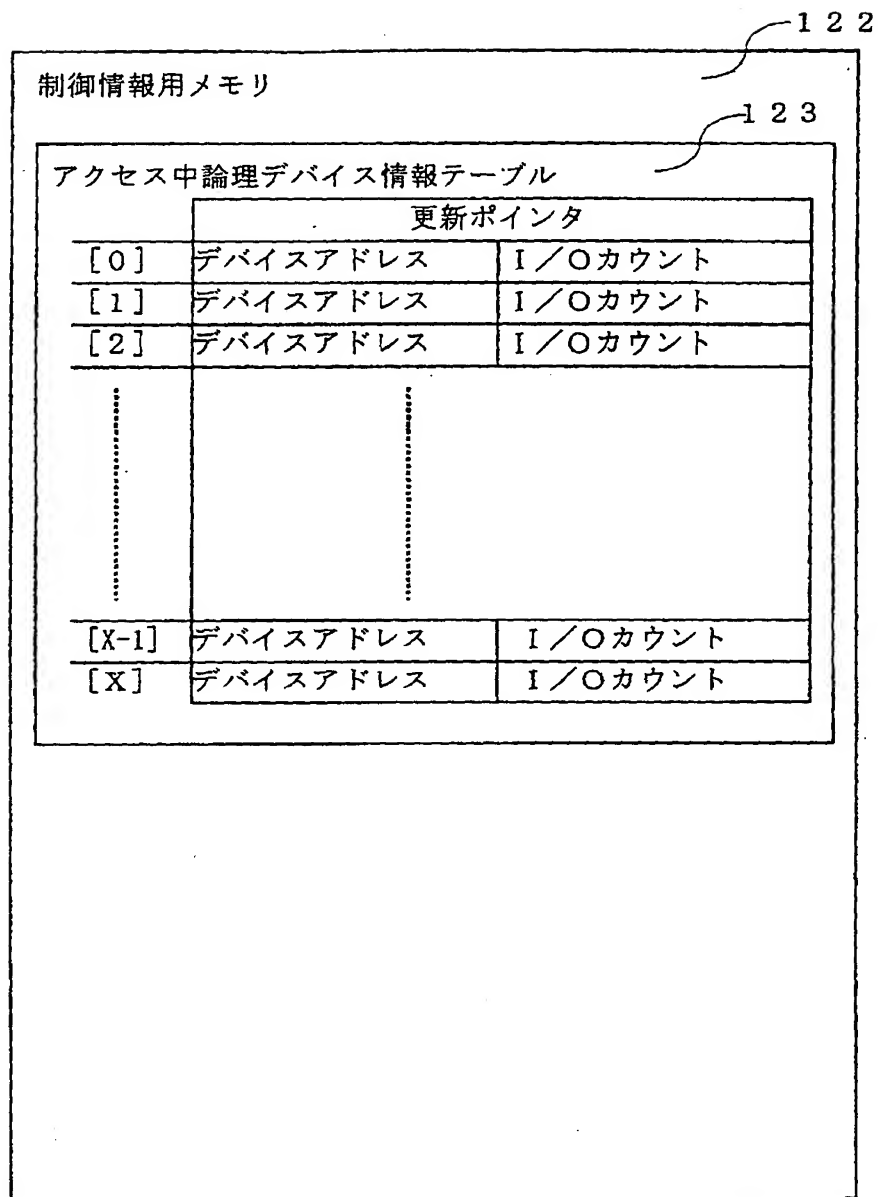
特2002-240719

図2

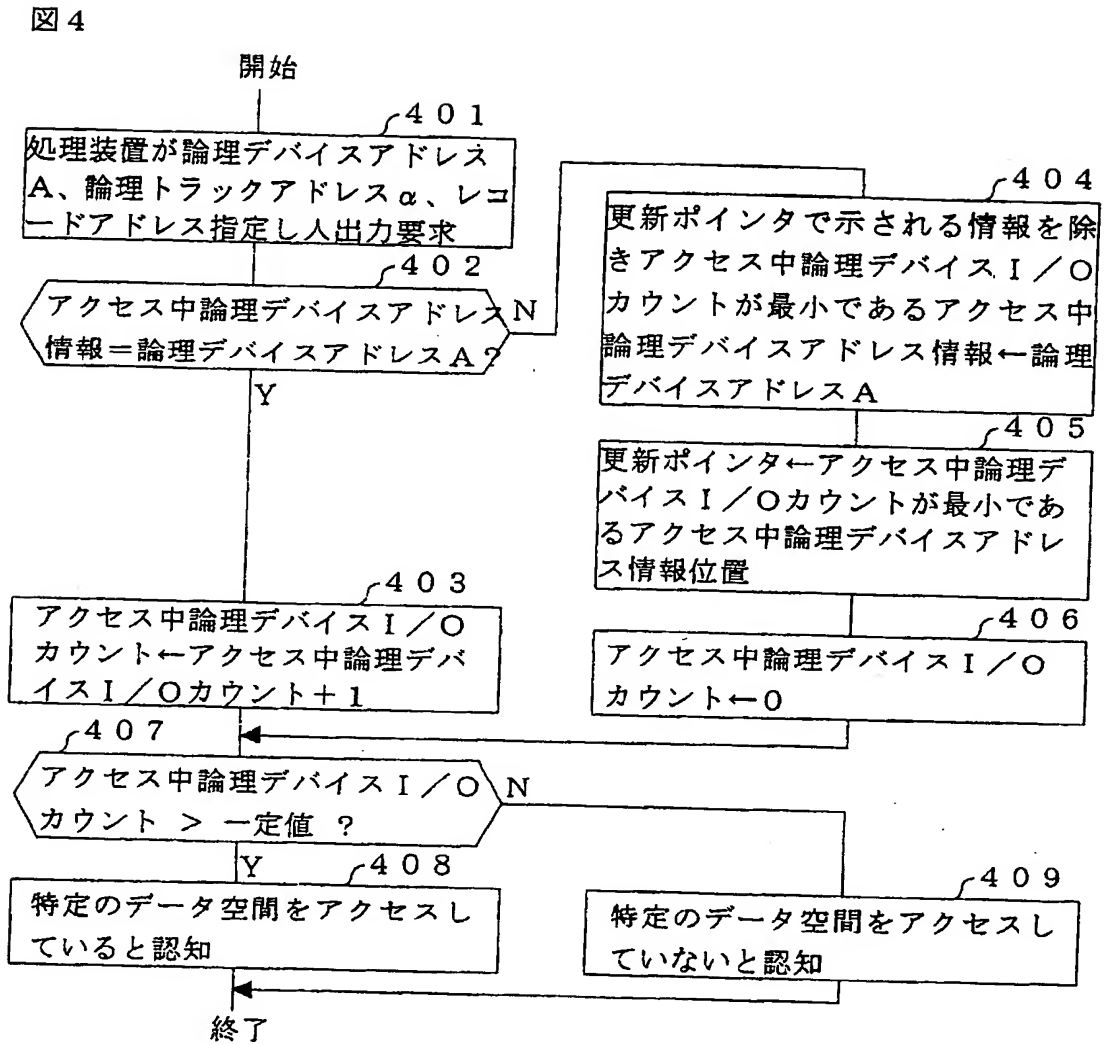


【図3】

図 3

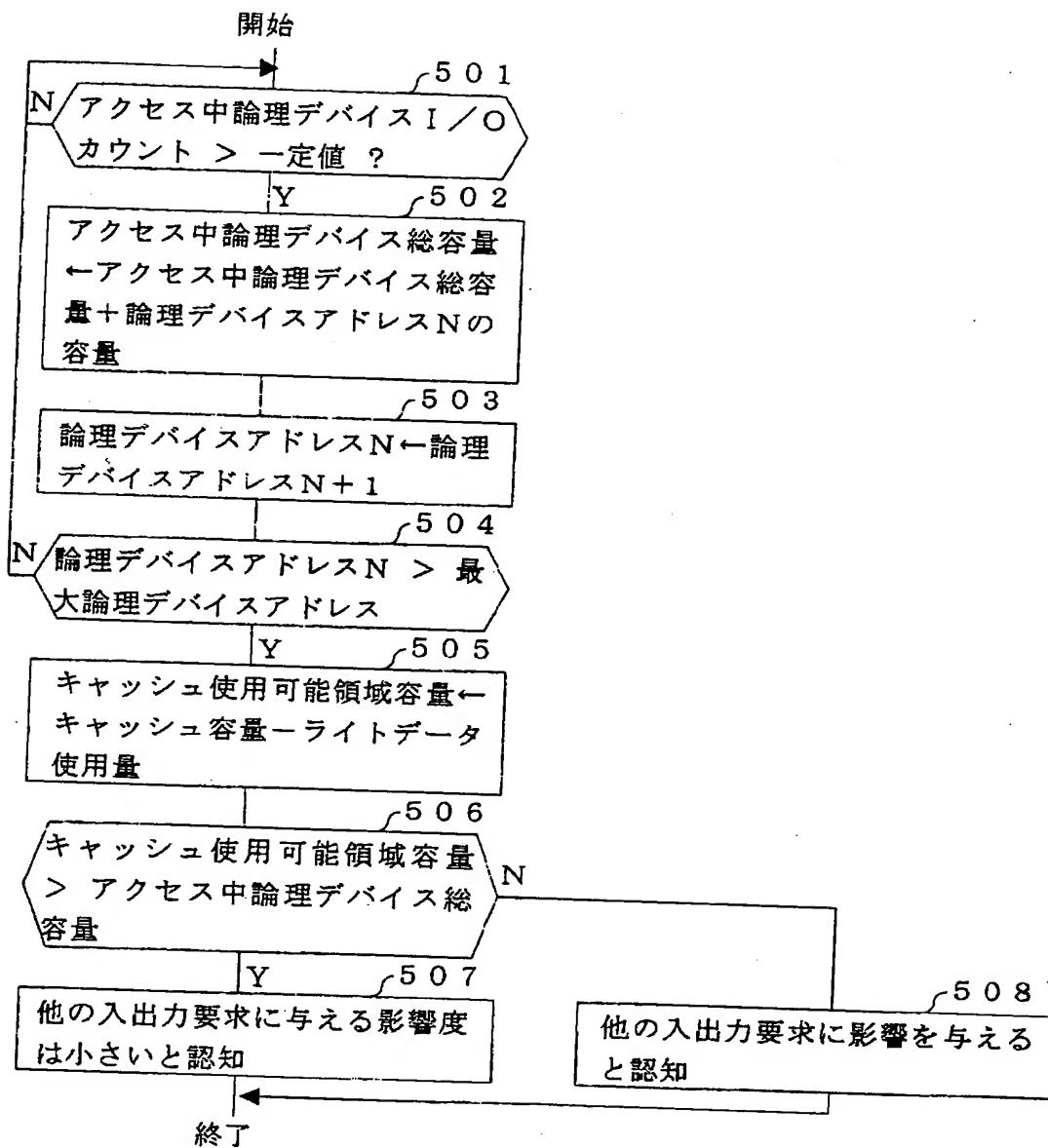


【図4】



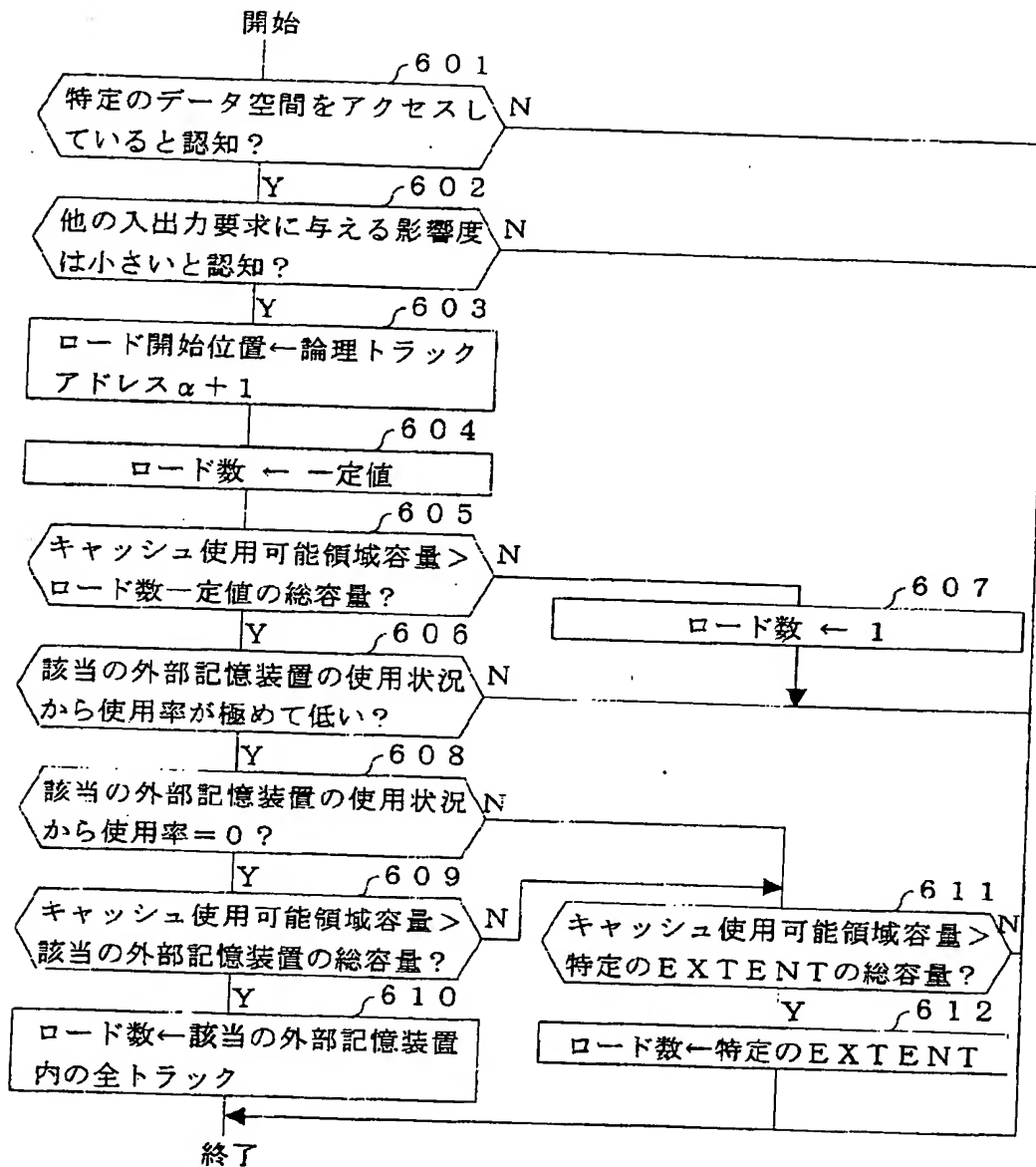
【図5】

図5

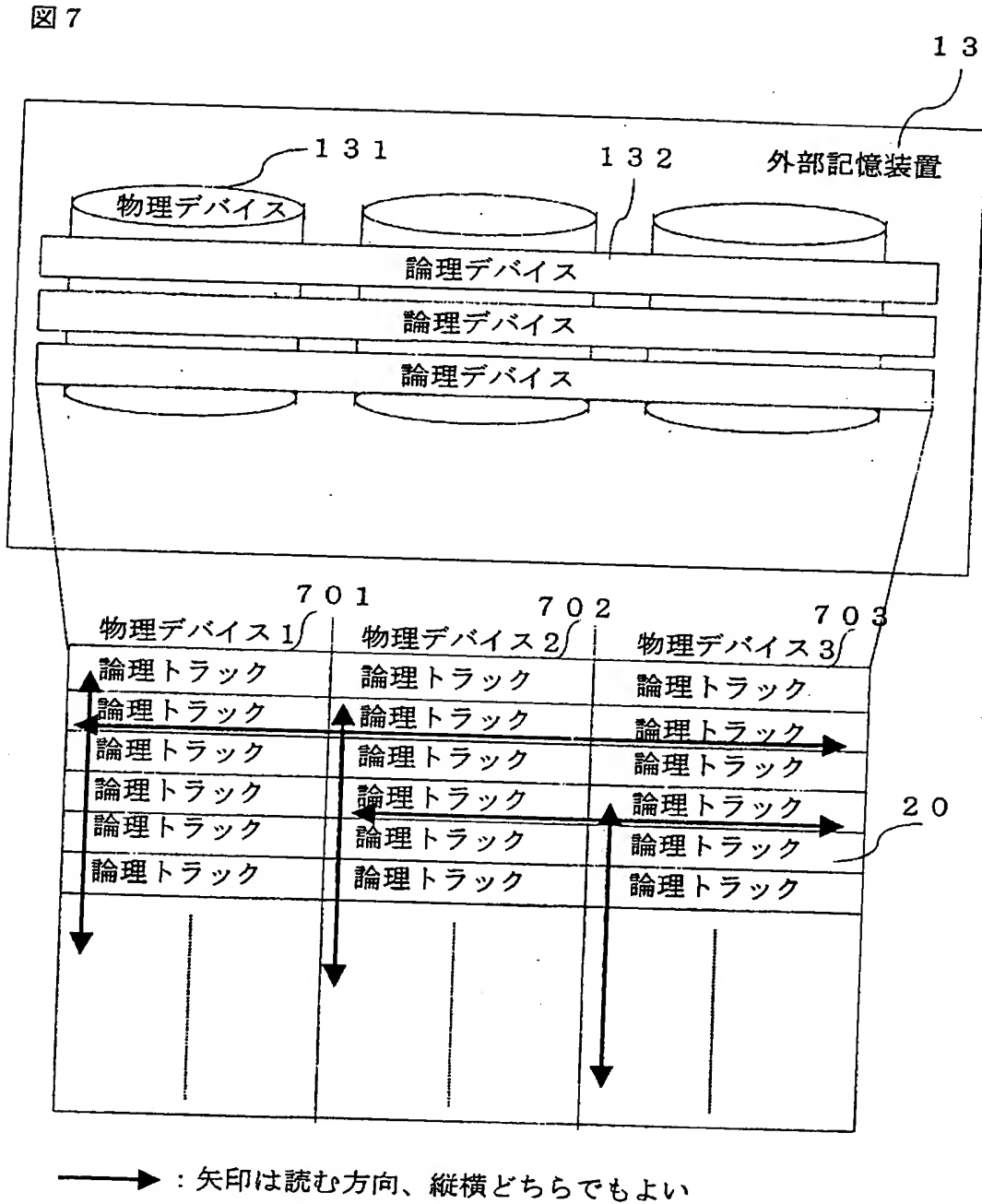


【図 6】

図 6



【図 7】



【書類名】 要約書

【要約】

【課題】 ランダムアクセスであっても即効性のある方式で高キャッシュヒット率を実現してレスポンスを向上させるデータ先読み制御を提供すること。

【解決手段】 現時点の入出力要求が、最近の入出力要求であってI/Oカウン
トが一定数を超える特定データベースにアクセスしていることを判断し、キャッ
シュメモリをある程度占有でき他の入出力要求に影響を与えることが無いと判断
した場合に、I/O要求されたブロック（論理トラック）よりも十分大きな複数
ブロックを含んでキャッシュメモリへロードする。最新の入出力要求のデータベ
ースを重視し、キャッシュメモリの使用率を最大限利用するものである。

【選択図】 図 6

出 願 人 履 歴 情 報

識別番号

[000005108]

1. 変更年月日

1990年 8月31日

[変更理由]

新規登録

住 所

東京都千代田区神田駿河台4丁目6番地

氏 名

株式会社日立製作所